

Europäisches Patentamt  
European Patent Office  
Office européen des brevets



(11) **EP 1 363 272 A1**

(12) **DEMANDE DE BREVET EUROPEEN**

(43) Date de publication:  
19.11.2003 Bulletin 2003/47

(51) Int Cl.7: **G10L 19/04, H04M 1/725**

(21) Numéro de dépôt: **03291085.3**

(22) Date de dépôt: **06.05.2003**

(84) Etats contractants désignés:  
**AT BE BG CH CY CZ DE DK EE ES FI FR GB GR  
HU IE IT LI LU MC NL PT RO SE SI SK TR**  
Etats d'extension désignés:  
**AL LT LV MK**

(30) Priorité: **16.05.2002 FR 0206035**

(71) Demandeur: **ALCATEL**  
**75008 Paris (FR)**

(72) Inventeurs:  
• **Bonnard, Pierre**  
**92150 Suresnes (FR)**

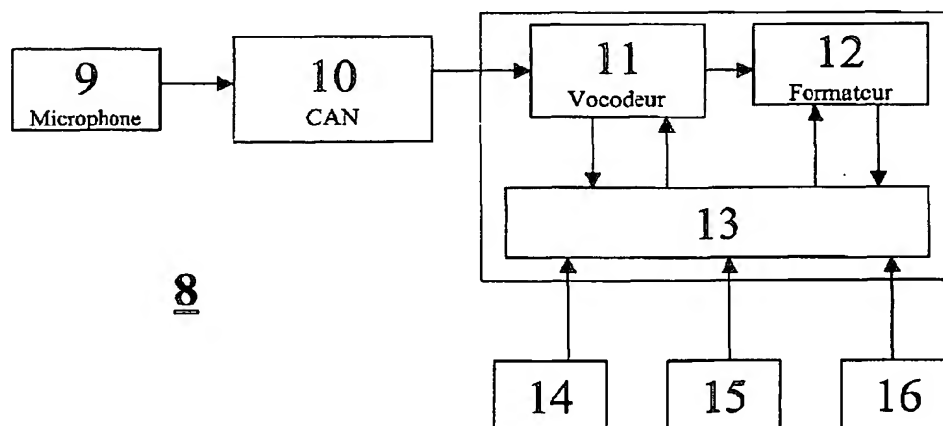
• **Bourmeyster, Ivan**  
**75015 Paris (FR)**  
• **Fourquin, Xavier**  
**92310 Sevres (FR)**  
• **Ladouce, Pierre**  
**95130 Franconville (FR)**

(74) Mandataire: **Korakis-Ménager, Sophie**  
**Compagnie Financière Alcatel**  
**Département de Propriété Industrielle,**  
**5, rue Noel Pons**  
**92734 Nanterre Cedex (FR)**

(54) **Terminal de télécommunication permettant de modifier la voix transmise lors d'une communication téléphonique**

(57) La présente invention concerne un terminal de télécommunication permettant de modifier la voix transmise lors d'une communication téléphonique, plus particulièrement adapté à un système de téléphonie mobile. A cet effet, le terminal (8) comporte des moyens (9) d'entrée par l'utilisateur dudit terminal d'un signal de parole analogique, des moyens (10) pour convertir ledit si-

gnal de parole analogique en un signal numérique, un vocodeur (11) pour exécuter un codage de source dudit signal de parole, ledit vocodeur (11) incluant des moyens pour extraire des paramètres de reconstruction dudit signal de parole. Le terminal est caractérisé en ce qu'il comporte des moyens (13) pour modifier les paramètres de reconstruction de sorte que la voix transmise associée au dit signal soit modifiée.



**Fig. 1**

## Description

[0001] La présente invention concerne un terminal de télécommunication permettant de modifier la voix transmise lors d'une communication téléphonique, plus particulièrement adapté à un système de téléphonie mobile.

[0002] Même si la transmission de la parole reste l'élément essentiel de la téléphonie mobile il n'en demeure pas moins que les fabricants cherchent à différencier leurs produits en offrant de nouveaux services attractifs et distrayants pour le consommateur. Les jeux, les services liés à la reconnaissance vocale ou la multiplication des sonneries d'appel en sont des exemples.

[0003] Ces nouveaux services impliquent souvent un coût supplémentaire sur le téléphone lié à l'ajout d'éléments logiciels ou matériels.

[0004] La présente invention vise à fournir un terminal de télécommunication offrant un service de modification de la voix transmise par l'utilisateur du terminal lors d'une communication téléphonique, ce service ayant un caractère attractif et distrayant et étant mis en oeuvre de façon simple et économique.

[0005] La présente invention propose à cet effet un terminal de télécommunication comportant des moyens d'entrée par l'utilisateur dudit terminal d'un signal de parole analogique, des moyens pour convertir ledit signal de parole analogique en un signal numérique, un vocodeur pour exécuter un codage de source dudit signal de parole, ledit vocodeur incluant des moyens pour extraire des paramètres de reconstruction dudit signal de parole, **caractérisé en ce que** ledit terminal comporte des moyens pour modifier lesdits paramètres de reconstruction de sorte que la voix transmise associée au dit signal soit modifiée.

[0006] Grâce à l'invention, la mise en place d'un service de modification de la voix transmise par un téléphone mobile se fait de façon simple et économique en utilisant un vocodeur déjà présent dans le téléphone pour le codage de la voix. Les paramètres de reconstruction sont représentatifs de la voix à transmettre et leur modification permet d'agir directement sur la sonorité de la voix transmise.

[0007] Avantageusement, lesdits moyens pour modifier les paramètres de reconstruction comportent des moyens pour modifier la fréquence fondamentale de la voix.

[0008] Selon un mode de réalisation, ledit codage de source du signal de parole est un codage RPE-LTP (Regular Pulse Excitation Long Term Prediction).

[0009] De manière avantageuse, lesdits moyens pour modifier les paramètres de reconstruction comportent des moyens pour modifier les paramètres de filtre court terme dudit vocodeur.

[0010] Les paramètres de filtre court terme sont par exemple les coefficients de réflexion, dits coefficients LAR (Log Area Ratio) déterminés pendant un codage du type RPE-LTP.

[0011] Selon une variante, lesdits moyens pour modifier les paramètres de reconstruction comportent des moyens pour modifier les paramètres de filtre long terme dudit vocodeur.

5 [0012] Un paramètre de filtre long terme est par exemple la fréquence fondamentale associée à un pitch déterminé pendant un codage du type RPE-LTP.

[0013] Avantageusement, lesdits moyens pour modifier les paramètres de reconstruction comportent des moyens pour modifier les paramètres associés à des grandeurs d'excitation dudit vocodeur.

10 [0014] Un paramètre associé à une grandeur d'excitation peut être par exemple un paramètre définissant le signal impulsionnel d'excitation déterminé pendant un codage RPE-LTP.

15 [0015] Selon un mode de réalisation, lesdits moyens pour modifier les paramètres de reconstruction comportent des moyens pour modifier lesdits paramètres en fonction de données codées représentatives d'un ensemble de notes définies par leurs fréquences et leurs durées respectives.

[0016] Avantageusement, le terminal comporte des moyens de sélection d'un type de modification de la voix à transmettre.

20 [0017] De manière avantageuse, le terminal comporte des moyens pour configurer manuellement les modifications desdits paramètres de reconstruction du signal de parole.

[0018] Selon un mode de réalisation, le terminal comporte des moyens pour exécuter lesdites modifications dans un mode de test.

[0019] De manière avantageuse, le terminal est un téléphone mobile.

25 [0020] D'autres caractéristiques et avantages de la présente invention apparaîtront dans la description suivante d'un mode de réalisation de l'invention donné à titre illustratif et nullement limitatif.

[0021] Dans les figures suivantes :

- 40 • La figure 1 représente schématiquement un terminal de télécommunication selon l'invention,
- La figure 2 représente schématiquement les moyens de codage et de modification de la voix utilisés dans un terminal de télécommunication selon l'invention, utilisant un codage RPE-LTP conforme au standard GSM 06.10.

[0022] Dans toutes les figures, les éléments communs portent les mêmes numéros de référence.

50 [0023] La figure 1 représente un terminal de télécommunication 8 selon l'invention tel qu'un téléphone mobile.

[0024] Le terminal 8 comporte un microphone 9, un convertisseur analogique numérique 10, un vocodeur 11, un formateur de données 12, des moyens 13 pour modifier des paramètres de reconstruction de la voix, des moyens 14 de sélection d'un type de modification de la voix à transmettre, des moyens 15 pour configurer

manuellement les modifications des paramètres de reconstruction du signal de parole et des moyens 16 pour exécuter des modifications dans un mode de test.

[0025] Le vocodeur 11, le formateur de données 12 et les moyens 13 pour modifier les paramètres de reconstruction de la voix vont être décrits ci dessous en relation avec la figure 2.

[0026] La figure 2 représente un vocodeur 11 pour exécuter un codage RPE-LTP conforme au standard GSM 06.10, un formateur 12 de flux de données et des moyens 13 de modifications de paramètres de reconstruction de la voix. Le vocodeur 11, le formateur 12 et les moyens 13 sont des moyens logiciels exécutés par un processeur programmable non représenté.

[0027] Le vocodeur 11 inclut des moyens 2 de détermination de coefficients de réflexion LAR, des moyens 3 de détermination d'une fréquence fondamentale de la voix et des moyens 4 de détermination d'un signal impulsif d'excitation, à partir d'un signal S échantillonné.

[0028] Le principe du codage RPE-LTP consiste à modéliser le signal de parole comme un filtre linéaire dont les paramètres évoluent dans le temps, ledit filtre linéaire étant excité par un signal impulsif. Dès lors, on comprend que l'objectif de ce codage est de déterminer le signal impulsif d'excitation, les différents paramètres définissant le filtre linéaire et la fréquence fondamentale de la voix à transmettre. Ce type de codage de la parole opère sur des trames de 20 ms sur la bande [300-3400 Hz]. Le filtre linéaire se décompose en deux parties : une première partie qui prend en compte les corrélations à courts termes du signal de parole et une deuxième partie qui prend en compte les corrélations à long terme du signal de parole. On parlera dans ce qui suit de filtre court terme et de filtre long terme.

[0029] Le signal S constitue une trame de parole de 20 ms échantillonnée à 8 kHz.

[0030] Les moyens 2 permettent de déterminer les coefficients de réflexion LAR du filtre court terme. On utilise pour cela l'algorithme de Durbin qui est appliqué au signal filtré afin d'en obtenir une prédiction linéaire. L'algorithme de Durbin génère dix coefficients de réflexion, dits coefficients LAR, compris entre -1 et 1 exprimés à l'échelle logarithmique LAR (Log Area Ratio).

[0031] Les moyens 3 de détermination d'une fréquence fondamentale de la voix utilisent un procédé de prédiction associé au filtre long terme sur des trames de 5 ms et permettent de reproduire la fréquence fondamentale associée à un pitch P compris entre 56 et 444 Hz et caractéristique du locuteur ainsi que le gain associé à cette fréquence.

[0032] Les moyens 4 permettent de déterminer le signal impulsif I d'excitation.

[0033] L'ensemble des données fournies par les moyens 2, 3 et 4 pour une trame de 20 ms est formaté dans le formateur 12 de flux de données qui va produire un signal de parole C codé.

[0034] Lorsque ces paramètres sont définis, on peut produire différents effets sur la voix en modifiant lesdits paramètres LAR, P et I. Les moyens de modification 13 permettent de modifier ces différents paramètres.

[0035] Ainsi, les moyens 13 de modification peuvent modifier uniquement la valeur du pitch associée à la fréquence fondamentale de façon à transformer une voix d'homme en une voix de femme. Pour cela, la valeur de pitch P de chaque trame est modifiée en une valeur Min ( $P + P_{\text{effet}}$ , Max (P)),  $P_{\text{effet}}$  étant une valeur prédéfinie, par exemple à 90 Hz, et Max (P) étant la valeur de pitch maximale acceptée par le système standard. En effet, une voix masculine a une valeur de pitch à environ 120 Hz alors qu'une voix féminine a une valeur de pitch à environ 210 Hz.

[0036] Les moyens 13 de modification peuvent procéder de façon similaire pour transformer une voix de femme en une voix d'homme. Pour cela, la valeur de pitch P de chaque trame est modifiée en une valeur Max ( $P - P_{\text{effet}}$ , Min (P)),  $P_{\text{effet}}$  étant une valeur prédéfinie à 90 Hz et Min (P) étant la valeur de pitch minimale acceptée par le système standard.

[0037] On peut également obtenir un effet de chuchotement lorsque les moyens 13 modifient le pitch et le gain. Pour cela, les moyens de modification 13 augmentent le pitch et diminuent le gain de sorte que la sortie du signal du filtre long terme soit substantiellement modifiée en diminuant l'effet des vibrations des cordes vocales. Les paramètres LAR restent inchangés.

[0038] On peut aussi obtenir un effet de voix robotisée lorsque les moyens 13 modifient les paramètres du filtre long terme. Pour cela, le pitch est fixé à une valeur relativement élevée. Comme pour l'effet de chuchotement, la prosodie, c'est à dire l'ensemble des éléments phoniques qui caractérisent le langage parlé, est complètement modifiée et transformée en un son monotone ayant une fréquence fondamentale unique et des paramètres LAR inchangés.

[0039] Une autre propriété remarquable des paramètres des codeurs de voix concerne le fait que les paramètres LAR représentent le trajet vocal et le trajet nasal liés à la production de parole, ces deux trajets pouvant être modélisés par des filtres de résonance du type filtre « tout pôle », chaque pôle représentant une fréquence de résonance. Dès lors, la production de la parole faisant intervenir le trajet vocal et le trajet nasal, les moyens 13 peuvent modifier ces paramètres de sorte que les voyelles résonnent comme celles d'une personne qui a le nez bouché.

[0040] Le trajet vocal et le trajet nasal peuvent chacun être modélisés par un filtre tout pôle mais la somme de ces deux filtres n'est pas tout pôle. Pour modéliser la combinaison de ces deux trajets sous la forme d'un filtre tout pôle, il faut construire un zéro en plaçant deux pôles cote à cote. Ce zéro représente la contribution de la cavité nasale sur la cavité vocale.

[0041] En rapprochant ces deux pôles cote à cote, les moyens de modification 13 diminuent la contribution de

la cavité nasale et donnent ainsi un effet « enrhumé » à la voix.

[0042] Les paramètres peuvent également être modifiés en fonction de données codées représentatives d'un ensemble de notes définies par leurs fréquences et leurs durées respectives. Ainsi la voix modifiée peut suivre une partition définie dans un format de codage définissant un ensemble de couples (fréquence, durée) et avoir ainsi un effet chanté. Une partition est fournie aux moyens 13 de modification des paramètres de reconstruction. Cette partition peut être sous un format de codage de musique MIDI, SMAF de Yamaha®, EMR R5 polyphonique, IrDA iMelody de l'irMC (Infrared Mobile Communications) ou un autre format de description vectoriel de la musique. Dans une étape de prétraitement, la partition est transcrite automatiquement, si nécessaire, pour que les fréquences des notes soient dans l'intervalle [56 Hz, 444 Hz].

[0043] Les notes sont interprétées en durée et en fréquence, à chaque note correspondant une date de début, une date de fin et une fréquence. Les trames de parole de 5 ms sont modifiées par les moyens 13 pour que le pitch de la parole soit égal à la fréquence de la note correspondante au même instant.

[0044] Les modifications des paramètres se font généralement via des échanges entre le formateur de flux de données et les moyens 13 pour modifier les paramètres. Toutefois, on peut également envisager des modifications des paramètres faites directement au niveau du vocodeur 11.

[0045] En figure 1, l'utilisateur a la possibilité de sélectionner le type d'effets souhaité sur la voix à transmettre via les moyens 14 de sélection d'un type de modification. Il peut également configurer manuellement les modifications des paramètres de reconstruction du signal de parole via les moyens 15. Ainsi, il peut par exemple influencer sur la valeur du pitch modifié à Min ( $P + P_{\text{effet}}$ , Max (P)) dans le cas d'une transformation de voix masculine en voix féminine en fixant lui-même la valeur de  $P_{\text{effet}}$ . Il peut également tester l'effet obtenu via les moyens 16 qui lui permettent d'exécuter le type de modification sélectionnée avec les paramètres tels qu'il les a fixés.

[0046] Lors d'une communication mobile, la voix de l'utilisateur est d'abord échantillonnée et numérisée par les moyens 9. Elle est ensuite codée par le vocodeur 11, lequel met en oeuvre l'étape 7 tel que décrite en relation avec la figure 1.

[0047] La voix codée est ensuite formatée par le formateur 12 sous forme de paramètres de reconstruction de la voix obtenus grâce au vocodeur 11.

[0048] Les moyens 13, commandés par les moyens 14, 15 et 16 vont alors modifier les paramètres formatés de reconstruction de la voix pour obtenir les effets recherchés par l'utilisateur. Les paramètres modifiés sont ensuite retransmis aux moyens 12 puis sont envoyés à l'interlocuteur via le réseau de télécommunication ou exécutés en mode test pour l'utilisateur.

[0049] On peut également envisager une modification des paramètres de reconstruction directement au niveau des moyens 11 par les moyens 13.

[0050] Bien entendu, l'invention n'est pas limitée au mode de réalisation qui vient d'être décrit.

[0051] Notamment, le mode de réalisation porte sur un codage de source du type RPE-LTP mais on peut envisager tout autre codage de source tel que par exemple l'ETS 300 726 GSM EFR (Enhanced Full Rate) et le 3GPP TS 26.071 AM (Adaptive Multirate).

## Revendications

1. Terminal (8) de télécommunication comportant des moyens (9) d'entrée par l'utilisateur dudit terminal d'un signal de parole analogique, des moyens (10) pour convertir ledit signal de parole analogique en un signal numérique (S), un vocodeur (11) pour exécuter un codage de source dudit signal de parole (S), ledit vocodeur (11) incluant des moyens (11) pour extraire des paramètres de reconstruction dudit signal de parole, **caractérisé en ce que** ledit terminal (8) comporte des moyens (13) pour modifier lesdits paramètres de reconstruction de sorte que la voix transmise associée au dit signal soit modifiée.
2. Terminal (8) selon la revendication précédente **caractérisé en ce que** lesdits moyens (13) pour modifier lesdits paramètres de reconstruction comportent des moyens pour modifier un paramètre représentatif de la fréquence fondamentale de la voix.
3. Terminal (8) selon l'une des revendications précédentes **caractérisé en ce que** ledit codage de source du signal de parole est un codage RPE-LTP.
4. Terminal (8) selon l'une des revendications précédentes **caractérisé en ce que** lesdits moyens (13) pour modifier lesdits paramètres de reconstruction comportent des moyens pour modifier les paramètres de filtre court terme dudit vocodeur.
5. Terminal (8) selon l'une des revendications précédentes **caractérisé en ce que** lesdits moyens (13) pour modifier lesdits paramètres de reconstruction comportent des moyens pour modifier les paramètres associés à des grandeurs d'excitation.
6. Terminal (8) selon l'une des revendications précédentes **caractérisé en ce que** lesdits moyens (13) pour modifier lesdits paramètres de reconstruction comportent des moyens pour modifier les paramètres de filtre long terme dudit vocodeur.
7. Terminal (8) selon l'une des revendications précédentes **caractérisé en ce que** lesdits moyens (13)

pour modifier lesdits paramètres de reconstruction comportent des moyens pour modifier lesdits paramètres en fonction de données codées représentatives d'un ensemble de notes définies par leurs fréquences et leurs durées respectives.

5

8. Terminal selon l'une des revendications précédentes comportant des moyens (14) de sélection d'un type de modification de la voix à transmettre.

10

9. Terminal selon l'une des revendications précédentes comportant des moyens (15) pour configurer manuellement les modifications desdits paramètres de reconstruction du signal de parole.

15

10. Terminal selon l'une des revendications précédentes comportant des moyens (16) pour exécuter lesdites modifications dans un mode de test.

11. Terminal selon l'une des revendications précédentes **caractérisé en ce que** ledit terminal est un téléphone mobile.

20

25

30

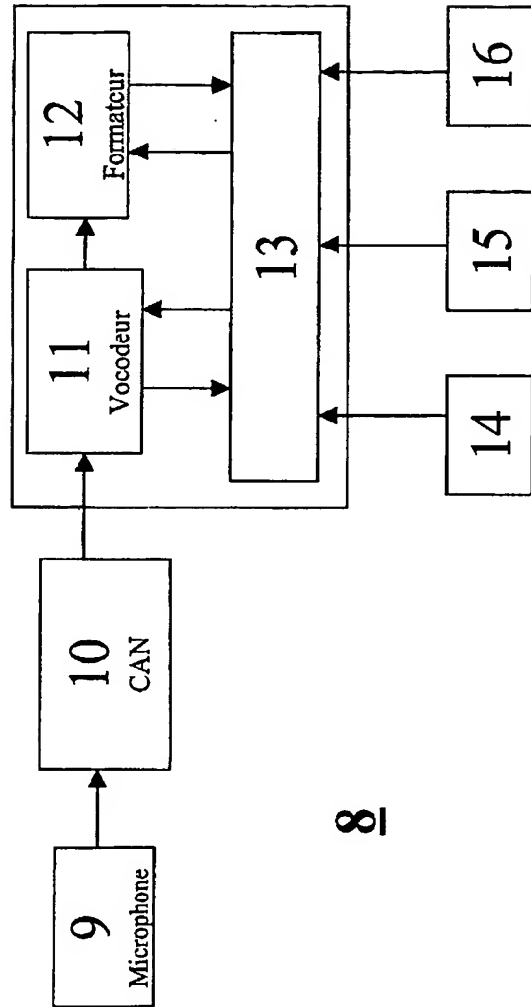
35

40

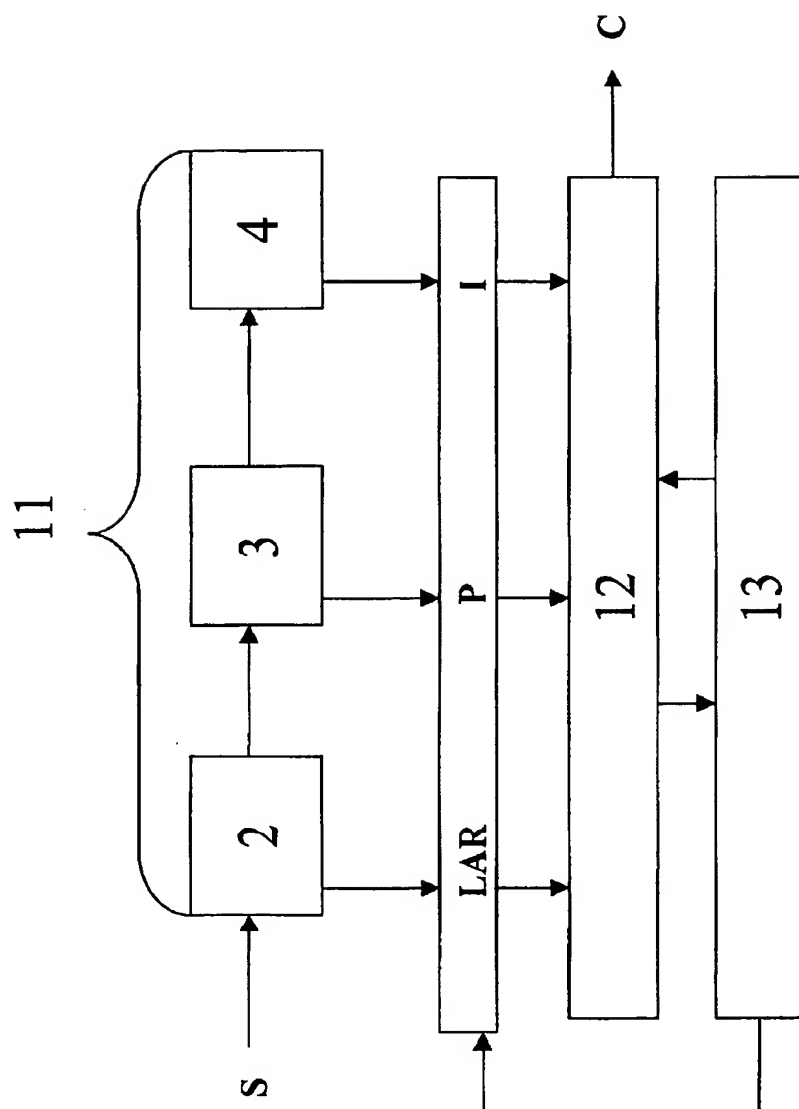
45

50

55



**Fig. 1**

**Fig. 2**



Office européen  
des brevets

# RAPPORT DE RECHERCHE EUROPEENNE

Numéro de la demande  
EP 03 29 1085

DOCUMENTS CONSIDERES COMME PERTINENTS			
Catégorie	Citation du document avec indication, en cas de besoin, des parties pertinentes	Revendication concernée	CLASSEMENT DE LA DEMANDE (Int.Cl.7)
X	US 5 692 101 A (HARTMAN ET AL) 25 novembre 1997 (1997-11-25) * le document en entier *	1,4-6,11	G10L19/04 H04M1/725
X	EP 1 126 439 A (LUCENT TECHNOLOGIES INC) 22 août 2001 (2001-08-22)	1,2,11	
A	* page 2, ligne 10 - ligne 56 * * page 3, ligne 20 - page 5, ligne 15; figures 1-10 *	3-10	
A	US 2001/006901 A1 (RYU BYUNG-SEOK) 5 juillet 2001 (2001-07-05) * le document en entier *	1,2,8,9, 11	
A	PATENT ABSTRACTS OF JAPAN vol. 2000, no. 01, 31 janvier 2000 (2000-01-31) & JP 11 289361 A (NEC CORP), 19 octobre 1999 (1999-10-19) * abrégé *	1,2,11	
A	US 4 823 380 A (KOHEN CHAIM ET AL) 18 avril 1989 (1989-04-18) * colonne 3, ligne 10 - colonne 5, ligne 7; figures 1,2 *	1,2	DOMAINES TECHNIQUES RECHERCHES (Int.Cl.7) G10L H04M
Le présent rapport a été établi pour toutes les revendications			
Lieu de la recherche LA HAYE		Date d'achèvement de la recherche 12 août 2003	Examineur Delangue, P
<p>CATEGORIE DES DOCUMENTS CITES</p> <p>X : particulièrement pertinent à lui seul Y : particulièrement pertinent en combinaison avec un autre document de la même catégorie A : arrière-plan technologique O : divulgation non-écrite P : document intercalaire</p> <p>T : théorie ou principe à la base de l'invention E : document de brevet antérieur, mais publié à la date de dépôt ou après cette date D : cité dans la demande L : cité pour d'autres raisons &amp; : membre de la même famille, document correspondant</p>			

EPO FORM 1503 (03.82) (P04C02)



**ANNEXE AU RAPPORT DE RECHERCHE EUROPEENNE  
RELATIF A LA DEMANDE DE BREVET EUROPEEN NO.**

EP 03 29 1085

La présente annexe indique les membres de la famille de brevets relatifs aux documents brevets cités dans le rapport de recherche européenne visé ci-dessus.  
Lesdits membres sont contenus au fichier informatique de l'Office européen des brevets à la date du  
Les renseignements fournis sont donnés à titre indicatif et n'engagent pas la responsabilité de l'Office européen des brevets.

12-08-2003

Document brevet cité au rapport de recherche		Date de publication	Membre(s) de la famille de brevet(s)	Date de publication
US 5692101	A	25-11-1997	AUCUN	
EP 1126439	A	22-08-2001	AU 1839001 A	16-08-2001
			BR 0100384 A	30-10-2001
			CA 2330985 A1	14-08-2001
			CN 1309521 A	22-08-2001
			DE 60001209 D1	20-02-2003
			EP 1126439 A2	22-08-2001
			JP 2001272998 A	05-10-2001
US 2001006901	A1	05-07-2001	KR 2001065803 A	11-07-2001
JP 11289361	A	19-10-1999	AUCUN	
US 4823380	A	18-04-1989	AUCUN	

EPO FORM P0460

Pour tout renseignement concernant cette annexe : voir Journal Officiel de l'Office européen des brevets, No.12/82